

# **Integrated Production Model (IPM), Release 5, Milestone 3/4A – Privacy Impact Assessment (PIA)**

**PIA Approval Date: July 1, 2009**

## **System Overview**

Integrated Production Model (IPM) supports the Business Operating Divisions (BODs) in the identification and selection of cases, measurement of noncompliance, and in the development of treatment strategies to deal with identified areas of noncompliance among individual and business filers. The data includes data from sources including the Master files (e.g. Individual Master File or IMF, Business Master file or BMF), the Return Transaction Files or RTFs, Transaction Data, third party data including information from SSA, HHS and Address information from Choicepoint, Employee Plans Information, Information Returns data, case information, large dollar cash transaction data, North American Industry Classification Code System (NAICS) data from the National Institute for Standards and Technology (NIST), State workforce agency data, demographics data and U.S. Census data.

## **Systems of Records Number (SORN):**

- Treasury/IRS 42.021 – Compliance Programs and Project Files

## **Data in the System**

### **1. Describe the information (data elements and fields) available in the system in the following categories:**

Data from Category A and D will be included in the system for Release 5. (See attached table for detailed list of Taxpayer Data)

#### **A. Taxpayer Data:**

- BMF Entities; Tax Modules; Entity Transactions; Tax Return Transactions; Tax Return Statuses
- Transcribed sections of Business Tax Returns: Forms 941 & 943
- Wage Related Documents from SSA
- Name Control; Date of Death; Birth Data; Citizenship Indicator
- Employee Plan Master File Entities and Tax Modules
- Information Return Documents; and Payer Documents
- Information Return Documents and Payer Documents
- IMF Entities; Tax Modules; Entity Transactions; Tax Return Transactions; Tax Return Statuses
- Examination Inventory and controls
- Transcribed sections of Individual Tax Returns: Form 1040
- Form 1099-MISC Income
- SS-8 Inventory and Case Development
- Earned Income Credit
- Calculated variables based on IMF tax returns as received from the Compliance Data Environment (CDE) System – debt to asset ratio, percent of expense(s) to total income

#### **B. Employee: There is no Employee data stored on the system.**

**C: Audit Trail Information:** Audit trails will be created using Oracle 10g auditing features. There are no individual end users that will access IPM. The audit tables within the DBS

are written to an operating system file each day and then the DBS tables are cleared. The file on the operating system is backed up to tape and retained for 7 years.

D. Other:

- Mapping Data that maps zip codes to Exam PODs were added to IPM at Release 4.
- Census Bureau; NAICS Codes; Industry Codes (OCC and SOC Codes)

**2. Describe/identify which data elements are obtained from files, databases, individuals, or any other sources.**

IPM Release 5 will include data from the following Sources:

A. IRS:

- Business Master File Database (BMF-DB2) – IRS Database
- Business Master File (BMF) – IRS Tier 1 Repository
- Business Return Transaction File (BRTF) – IRS Tier 1 Repository
- Combined Annual Wage Reporting (CAWR) – IRS Tier 1 Repository of wage related documents from SSA
- Compliance Data Environment (CDE) – IMF derived variables calculated from IMF transactions
- Data Master – 1 (DM-1) – IRS Tier 1 Repository
- Employee Plan Master File (EPMF-DB2) – IRS Database
- Information Returns Master File (IRMF) – IRS Tier 1 Repository
- Payer Master File (PMF) – IRS Tier 1 Repository
- Mapping Data that maps zip codes to Exam PODs.
- Individual Master File (IMF) – IRS Tier 1 Repository
- Individual Return Transaction File (IRTF) – IRS Tier 1 Repository
- Audit Information Management System (AIMS) - IRS Database
- Employment Tax Examination Program (ETEP) - IRS Database
- SS8ICP – IRS Database
- IPM Date Inputs (AIR File) – Flat File from BMFCCNIP

B. Taxpayer: No data is received directly from Taxpayers, as all taxpayer information is secured from an IRS source as listed in A. above.

C. Employee: No data is received directly from Employees

D. Other Federal Agency: Census Bureau – Manual .xls download from [http: site](http://site). IPM will not connect to this source.

E. State and Local Agencies: No data is received directly from State and Local Agencies.

F. Other third party sources: No data is received from other third party sources.

**3. Is each data item required for the business purpose of the system? Explain.**

Yes. The system is designed to serve as a common data repository for a number of client projects and system. All of the information identified as being required for IPM was derived by reviewing the client applications needs.

#### **4. How will each data item be verified for accuracy, timeliness, and completeness?**

Data contained in IPM comes from IRS authoritative sources which IPM relies on for accuracy. Data record element fields which are input to the Data Preparation component have already been extensively validated by their respective components. For example, delinquency data obtained from Business Master File has already been validated by that system.

Run-to-run balancing controls will be maintained by all programs which extract data for the IPM. The Log Accounting Reporting System (LARS) will be utilized to maintain control totals.

All Data Loads will employ the following Data Validation Rules:

- BMF and EPMF data is loaded as is from the master files on DB2. The exception to this is that any data that is acceptable on Tier I (DB2), but not acceptable on Tier II (Oracle), will be nulled.

For all other sources:

- When a value that is to be loaded into a DATE formatted attribute is found to be invalid, the value of the attribute is set to "null". The SSA\_DM1\_REF table is an exception. Dates from SSA are loaded as received since some are known to be invalid.
- When a value that is to be loaded into a NUMERICAL formatted attribute is found to be invalid, the value of the attribute is set to "null".
- When a new attribute is added to a table, the value of that attribute will be null in any rows that were already on the database.
- Right Trim (RTRIM) is used on all Name and Address attributes as well as the Condition Code Field and City attributes to conserve space on these variable length fields. Spaces to the right of the last significant character in these attributes are set to null.
- Money fields are represented as taken from the master files, dollars only or dollars and cents depending on the legacy field definition.

#### **5. Is there another source for the data? Explain how that source is or is not used.**

IPM extracts data from various internal authoritative sources into a centralized location. These authoritative sources, as listed in Question 2, are the original sources of the data contained in IPM; however, there is no other source that centralizes this data into one database.

#### **6. Generally, how will data be retrieved by the user?**

Individual users will not have access to the IPM to retrieve data. The system's design is that it will be a data repository that will be accessed by client applications through a direct link. All client applications will connect to IPM via an application to DB connection. There is no direct IPM User Interface. The design of the using application to IPM DB interface is the responsibility of the using application

#### **7. Is the data retrievable by a personal identifier such as name, SSN, or other unique identifier?**

Yes, data is retrievable by TIN: SSN; EIN. The client projects will build queries that will be run against IPM to retrieve data that fits the queries. These queries can contain personal identifiers.

## **Access to the Data**

### **8. Who will have access to the data in the system (Users, Managers, System Administrators, Developers, Others)?**

System and Database Administrators will have access to maintain the system. Developers will have access to the development system and read access to the production domain to check tables. Testers will have access to the test development system. Database Administrators, System Administrators, Developers and Project Account Managers will be granted access via OL5081 process.

As is stated above, the primary access to the system will be approved client applications. Users will gain access to the IPM data by way of these client applications designed to grant access.

### **9. How is access to the data by a user determined and by whom?**

IPM will have no individual end users with access to the data. Individual user data access will be determined at the client application level, with access controlled by the user's manager via the OL5081

### **10. Do other IRS systems provide, receive, or share data in the system? If YES, list the system(s) and describe which data is shared.**

The IPM will receive data from a number of systems. For release 5, the data sources will be BMF, BRTF, CDE, PMF, IRMF, EPMF, CAWR, DM-1, IMF, AIMS, IRTF, ETEP, SS8ICP and IPM Data Inputs. Each of these items is detailed in Question 2 above.

The IPM is designed to be a data repository. IPM Release 4.0 included data to accommodate client applications. Release 5.0 is for the purpose of expanding the data that is included in the IPM. Future client connections to IPM will be submitted as a change request (CR). The client CR will then be reviewed for approval and security impact by the IPM Configuration Control Board (CCB), which will also include members from Cybersecurity, ITSAE, and Systems Architecture and Engineering. Access to the IPM system can not occur until IPM is in production. At release 4.0 the Client Applications SWETRS and RBSS were provided access to IPM.

### **11. Have the IRS systems described in Item 10 received an approved Security Certification and Privacy Impact Assessment?**

No Client Applications will be allowed access to the IPM without an approved Security Certification and Accreditation (C&A). Each client application must request access through the IPM Configuration Control Board prior to gaining access to IPM. The connecting application must have a C&A that permits connection to IPM. The connection method must also be one that is currently supported by IPM. Potential clients will also need to submit certain documentation such as the IPM ICD, MOU, ATO, etc. The client CR will then be reviewed for approval by the IPM Configuration Control Board (CCB), which will also include members from Cybersecurity, ITSAE, and Systems Architecture and Engineering.

Prior to Release 5.0, client applications BMFCCNIP, RBSS and SWETRS gained access to the IPM.

### **12. Will other agencies provide, receive, or share data in any form with this system?**

IPM will receive data from wage related documents from SSA, as listed in 1A. above, however, the source of this data for IPM is CAWR, which is an internal IRS System. IPM will also receive Census Data, as described in 2f above. This data will be secured via a manual download from the Census Bureau website.

## **Administrative Controls of Data**

### **13. What are the procedures for eliminating the data at the end of the retention period?**

IPM does not contain original documents or data and is not the authoritative source for data. The data loaded into IPM from the various data sources listed in this questionnaire will be updated with current information from weekly, monthly or annual loads.

W-4 paper case files and National Computer Center (NCC) magnetic tape records are destroyed two years after the cases have been inactive.

Records associated with the examination of returns are disposed in accordance with the established Retention Schedule 1(15)59.22, Records Control Schedule. Records not associated with the examination of returns will be destroyed at the completion of the program or project, or in accordance with Records Disposition Handbooks, IRM 1.15.2.1 through IRM 1.15.2.31.

**14. Will this system use technology in a new way?**

No. The IPM will use IRS Enterprise Architecture approved systems and software to store data compliance related data in a format that is easily accessed by the client applications

**15. Will this system be used to identify or locate individuals or groups? If so, describe the business purpose for this capability.**

Yes. The Client Applications will access data from the IPM that will be matched against their business rules that will include case selection. The business purpose for this capability is to ensure that taxpayers are compliant in their tax obligations. The system will be used to address tax gap issues.

**16. Will this system provide the capability to monitor individuals or groups? If yes, describe the business purpose for this capability and the controls established to prevent unauthorized monitoring.**

The IPM will be designed to allow compliance functions to identify non compliant taxpayers. The system is not designed to monitor individuals or groups except for those taxpayers or groups of taxpayers who are non compliant.

**17. Can use of the system allow IRS to treat taxpayers, employees, or others, differently? Explain.**

The IPM will be designed to allow compliance functions to identify non compliant taxpayers. The system is not designed to nor will it contain information on Race, Gender, Sexual Preference, or any other issue not related to tax compliance.

**18. Does the system ensure "due process" by allowing affected parties to respond to any negative determination, prior to final action?**

The system does not make determinations negative or otherwise. These actions are in the purview of the client applications and ensuring "due process" is the responsibility of these client applications.

**19. If the system is web-based, does it use persistent cookies or other tracking devices to identify web visitors?**

The system is not web based.

[View other PIAs on IRS.gov](#)